Ancient gene duplication and domain shuffling in the animal cyclic nucleotide phosphodiesterase family

Mitsumasa Koyanagi, Hiroshi Suga, Daisuke Hoshiyama, Kanako Ono, Naoyuki Iwabe, Kei-ichi Kuma, Takashi Miyata*

Department of Biophysics, Faculty of Science, Kyoto University, Kyoto 606, Japan

Received 20 August 1998

Abstract The animal cyclic nucleotide phosphodiesterases (PDEs) comprise at least seven subtypes, PDE1–7, which differ from each other in domain organization and primary function, and they diverged from an ancestral gene by gene duplication and domain shuffling during animal evolution. To obtain rough estimates for the divergence times of these subtypes, cloning of PDE cDNAs from *Ephydatia fluviatilis* (freshwater sponge) by RT-PCR was carried out. We obtained four cDNAs, EFPDE1, EFPDE2, EFPDE3, and EFPDE4, which are possibly homologs of the vertebrate PDE1, PDE2, PDE3, and PDE4, respectively, judging from the sequence similarity, domain organization, and branching pattern in the phylogenetic tree. The phylogenetic tree of the PDE family revealed that most gene duplications and domain shufflings that gave rise to different subtypes had been completed in the early evolution of animals before the separation of sponges and eumetazoans.

Key words: Phosphodiesterase; Sponge; Gene duplication; Domain shuffling; Phylogenetic tree; Gene diversity; Evolution

1. Introduction

Animals have evolved many gene families, each of which diverged from one or a few ancestral genes by gene duplication during animal evolution [1]. The animal cyclic nucleotide phosphodiesterase (PDE) family comprises at least seven subtypes or subfamilies, PDE1–7, which are distinguished from each other by domain organization and primary function [2,3]. Because yeast PDE consists of the catalytic domain alone [4], it is likely that the seven subtypes of animal PDE diverged from an ancestral gene having a single domain shared with fungal PDE by gene duplication and domain shuffling (hereafter we will refer to gene duplication that gave rise to different subtypes as subtype duplication). In addition there exist several isoforms in each subtype. The structure and function of these isoforms are virtually identical in the same subtype, but differ in tissue distribution (tissue specific genes). These isoforms were generated by gene duplication during evolution of vertebrates and arthropods (isoform duplication) [1,5].

To discover whether the number of genes involved in cell-cell communication specific to animals increased dramatically in concert with the Cambrian explosion, the burst of animal diversification at the Precambrian-Cambrian boundary [6], we recently cloned and sequenced the G protein α subunit (Gα) cDNAs and the protein tyrosine kinase (PTK) cDNAs from the freshwater sponge and the hydra. The phylogenetic trees of these families revealed that most subtype duplications had been completed in the early evolution of animals before the parazoan (sponge)-eumetazoan split about 940 million years (Myrs) ago, long before the Cambrian explosion (Suga et al., submitted). These results suggest no direct relationship between functional diversification of genes by subtype duplication and the Cambrian explosion. However, there is still a possibility of a link between the Cambrian explosion and gene diversification by domain shuffling, because the PTK tree was inferred from a comparison of the kinase domain sequences alone.

By cloning and sequencing sponge PDE cDNAs homologous to vertebrate PDEs, and a phylogenetic analysis of the PDE family, we report here that the domain shufflings are also very old, going back to before the parazoan-eumetazoan split.

2. Materials and methods

2.1. Isolation and sequencing of sponge PDE cDNAs

Poly(A)$^+$ RNA of *Ephydatia fluviatilis* (freshwater sponge) was extracted from the cells hatched from the gemmules [7] using the QuickPrep mRNA isolation kit (Pharmacia). *E. fluviatilis* cDNAs were synthesized by reverse transcriptase (SuperScript II, Gibco) using oligo(dT) primer. These cDNAs were used as templates for PCR amplification with Expand High Fidelity PCR System (Boehringer Mannheim). The sense and antisense degenerate primers were designed from conserved amino acid residues within the PDE catalytic domain as follows: (1) 5′-GTGGATCCCA(T/C)AA(T/C)IIIINNC-A(T/C)GC-3′, corresponding to the amino acid sequence HNXXHA, and (2) 5′-GCGGATCCGTN(T/C)TNGA(G/A)AA(T/C)- CA(T/C)-CA-3′, corresponding to VLENHH for sense primers; (3) 5′-CA-GAATTC(G/C)(G/T)(G/A)TG(G/A)TCI(A/T)N(G/A)TC(G/A)TG-3′, corresponding to HD(I/L/V/Y)DH(R/T/P), and (4) 5′-GTGA-ATTCCC(T/C)TGIII(G/A/C)(A/T/C)A(G/A)AA(T/C)TC-3′, corresponding to EF(F/W/Y)XQG for antisense primers. Each primer contains *Bam*HI or *Eco*RI restriction sites at the 5′ end (underlined). The first round of PCR amplification with primers 1 and 4 was conducted as follows: 2-min denaturation step at 94°C; then 5 cycles of 94°C (1 min), 46°C or 48°C (2 min), and 72°C (5 min); followed by 30 cycles of 94°C (30 s), 60°C (1 min), and 72°C (2 min); and finally 1 cycle of 60°C (5 min) and 72°C (10 min). Second rounds of PCR with nested primers 1 and 3 or 2 and 4 were carried out with primary amplification products. The PCR products were separated in a 1.5% agarose gel containing ethidium bromide. Products of expected size were isolated as gel slices, purified using GeneCleanII (Bio 101) and cloned into the pT7Blue vector (Novagen). Then, *Escherichia coli* strain DH5α (Toyobo) was transformed with ligated vector. More than

*Corresponding author.

Abbreviations: PDE, phosphodiesterase; Gα, G protein α subunit; PTK, protein tyrosine kinase; Myr, million year

The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases with accession numbers AB017021–AB017024.

PII: S0014-5793(98)01150-8

a)

```
                                                              o    o    o          o       o               o    o    o   - -o                            o

 1. sponge EFPDE1          LFAEQMDSWNFNI--LHAEQLCS---NAPLRHIGHELFRKHDLTKKYKILPSCLDAFLAKVEAGYQLHGNPYHNSTHAADVLQTVHYLISST( 1)
 2. nematode T04D3.3       LMKE-VCTWSFSP--FQLNEVSE---GHALKYVGFELFNRYGFMDRFKVPLTALENYLSALEVGYSKHNNPYHNVVHAADVTQSSHFMLSQT( 1)
 3. human PDE1A            TLKD-VDKWSFDV--FALNEASG---EHSLKFMIYELFTRYDLINRFKIPVSCLITFAEALEVGYSKYKNPYHNLIHAADVTQTVHYIMLHT( 1)
 4. human PDE1B            CLKN-LDLWCFDV--FSLNQAAD---DHALRTIVFELLTRHNLISRFKIPTVFLMSFLDALETGYGKYKNPYHNQIHAADVTQTVHCFLLRT( 1)
 5. human PDE1C            ALKD-VDKWSFDV--FSLNEASG---DHALKFIFYELLTRYDLISRFKIPISALVSFVEALEVGYSKHKNPYHNLMHAADVTQTVHYLLYKT( 1)
 6. sponge EFPDE2          PLPP-VSSFHEHMCKLTFSPLDIK--NSADEIAVMSMMNEMDLINKFQLHPDILARFVIMVKKGY-R-DPPYHNWMHAFSVAHFLYALYCCS( 1)
 7. human PDE2A            GIQP-VAAIDSNFASFTYTPRSLP--EDDTSMAILSMLQDMNFINNYKIDCPTLARFCLMVKKGY-R-DPPYHNWMHAFSVSHFCYLLYKNL( 1)
 8. sponge EFPDE3          VLDE-VANWDFPI--FELREKSG---DHILSQMAYCLFNEAGLMETFKIPHRKFINFFRALEAHY--GNNVYHNKVHAADVLHATSYLLFEI(26)
 9. human PDE3A            IMEQ-LNTWNFPI--FDLVENIGRKCGRILSQVSYRLFEDMGLFEAFKIPIREFMNYFHALEIGY-R-DIPYHNRIHATDVLHAVWYLTTQP(45)
10. human PDE3B            LIEK-MSNWNFPI--FELVEKMGEKSGRILSQVMYTLFQDTGLLEIFKIPTQQFMNYFRALENGY-R-DIPYHNRIHATDVLHAVWYLTTRP(45)
11. sponge EFPDE4          EFFK-QEAWDMDI--FLLDTVSG---NSPLVAAAYHVFKARDLFSEMKIVPSVFLNFISTIESKY-H-SNPYHNALHAADVLLATNHLLKAK( 1)
12. nematode R153.1        HMQR-LDDWGPDV--FKIDELSK---NHSLTVVTFSLLRQRNLFKTFEIHQSTLVTYLLNLEHHY-R-NNHYHNFIHAADVAQSMHVLLMSP( 1)
13. Drosophila dunce       LLGE-LDTWGIQI--FSIGEFSV---NRPLTCVAYTIFQSRELLTSLMIPPKTFLNFMSTLEDHY-VKDNPFHNSLHAADVTQSTNVLLNTP( 1)
14. human PDE4A            ELEN-LNKWGLNI--FCVSDYAG---GRSLTCIMYMIFQERDLLKKFRIPVDTMVTYMLTLEDHY-HADVAYHNSLHAADVLQSTHVLLATP( 1)
15. human PDE4B            ELED-LNKWGLNI--FNVAGYSH---NRPLTCIMYAIFQERDLLKTFRISSDTFITYMMTLEDHY-HSDVAYHNSLHAADVAQSTHVLLSTP( 1)
16. human PDE4C            ELED-TNKWGLDV--FKVAELSG---NQPLTAIIFSIFQERDLLKTFQIPADTLATYLLMLEGHY-HANVAYHNSLHAADVAQSTHVLLATP( 1)
17. human PDE4D            ELED-VNKWGLHV--FRIAELSG---NRPLTVIMHTIFQERDLLKTFKIPVDTLITYLMTLEDHY-HADVAYHNNIHAADVVQSTHVLLSTP( 1)
18. bovine PDE5A           VVPS-AQTLKITD--FSFSDFELS--DLETALCTIRMFTDLNLVQNFQMKHEVLCKWILSVKKNY-RKNVAYHNWRHAFNTAQCMFAALKAG( 1)
19. human PDE6A            ELPD-ADKYEINK--FHFSDLPLT--ELELVKCGIQMYYELKVVDKFHIPQEALVRFMYSLSKGY-R-KITYHNWRHGFNVGQTMFSLLVTG( 1)
20. human PDE6B            ELPG-PTTFDIYE--FHFSDLECT--ELDLVKCGIQMYYELGVVRKFQIPQEVLVRFLFSISKGY-R-RITYHNWRHGFNVAQTMFTLLMTG( 1)
21. chicken PDE6C          ELPD-PKDLELYE--FRFSDFPVT--EHGLITCGIRLFFEINVVEKFKVPAEVLTRWMYTVRKGY-R-DITYHNWRHGFNVGQTMFTLLMTG( 1)
22. human PDE6C            DLPD-PRSAELYE--FRFSDFPLT--EHGLIKCGIRLFFEINVVEKFKVPVEVLTRWMYTVRKGY-R-AVTYHNWRHGFNVGQTMFTLLMTG( 1)
23. human PDE7A            MLEK-VGNWNFDI--FLFDRLTN---GNSLVSLTFHLFSLHGLIEYFHLDMMKLRRFLVMIQEDY-HSQNPYHNAVHAADVTQAMHCYLKEP( 1)
24. Saccharomyces PDE2     YWNI-LSTWDFCA--LSLS-------TQELIWCGFTLIKKLSKDAKVLIADNKLLLLLFTLESSY-HQVNKFHNFRHAIDVMQATWR-LCTY( 1)
                                                                     o          o          o   *      o**  *o

 1. LEHWISSLEVLAILLAATVHDLEHTGTTNTFHIQSYTDYALLYN-DRAVLENFHLNRAFTLL-KEDE-LNILKNLKLEDFRLVRSLVVEMVLATDMSSHSDQLKHMRQ( 9)QIEKG
 2. LANSLGDLELLAVLFGALIHDYEHTGHTNNFHIQSQSQFAMLYN-DRSVLENHHVSSCFRLM-KEDD-KNILTHLTRDEYKELRNMVIEIVLATDMSTHFMQIKTMKS( 4)PEGID
 3. IMHWLTELEILAMVFAAAIHDYEHTGTTNNFHIQTRSDVAILYN-DRSVLENHHVSAAYRLM-QEEE-MNILINLSKDDWRDLRNLVIEMVLSTDMSGHFQQIKNIRN( 4)PEGID
 4. MVHCLSEIELLAIIFAAAIHDYEHTGTTNSFHIQTKSECAIVYN-DRSVLENHHISSVFRLM-QDDE-MNIFINLTKDEFVELRALVIEMVLATDMSCHFQQVKTMKT( 4)LERID
 5. VANWLTELEIFAIIFSAAIHDYEHTGTTNNFHIQTRSDPAILYN-DRSVLENHHLSAAYRLL-QDDEEMNILINLSKDDWREFRTLVIEMVMATDMSCHFQQIKAMKT( 4)PEAIE
 6. KLSCLDDLEVLALFVSCLCHDIDHRGTNNAFQVCSNSTLACLYSSEGSVMERHHLAQTLCIL-NSPG-CNIFENLSDSDYRTTIQLIQDNILDTDIASHLKKLKHIKQ( 8)SNPEH
 7. LTNYLEDIEIFALFISCMCHDLDHRGTNNSFQVASKSVLAALYSSEGSVMERHHFAQAIAIL-NTHG-CNIFDHFSRKDYQRMLDLMRDIILATDLAHHLRIFKDLQK( 8)NNKQH
 8. VASAFSILEVFATYMAAAVHDFDHPGRTNAFLVATKSPLAILYN-DRAVLENHHCAASWDLLTSVPE-HNFLGDMEEAELKRFRFLLVEAVLATDLKMHFDLLSDFRA(12)SSEVD
 9. LSGNIPALELMALYVAAAMHDYDHPGRTNAFLVATSAPQAVLYN-DRSVLENHHAAAAWNLFMSRPE-YNFLINLDHVEFKHFRFLVIEAILATDLKKHFDFVAKFNG(10)TNEND
10. LSSNIPALELMALYVAAAMHDYDHPGRTNAFLVATNAPQAVLYN-DRSVLENHHAASAWNLYLSRPE-YNFLLHLDHVEFKRFRFLVIEAILATDLKKHFDFLAEFNA(12)SNEND
11. LEPIFTQLEVFAALVAAIIHDVDHPGRNNQFLVSTEDPLAILYN-DESVLENHHLALAFELM-KDPT-CNFLESTCKSQKQAFRKLIIDMVLATDMSRHFKYLGQLKT(15)DKYHD
12. LTEVFTDLEVLAAIFAGAVHDVDHPGFTNQYLINSNNELAIMYN-DESVLEQHHLAVAFKLL-QDSN-CDFLANLSRKQRLQFRKIVIDMVLATDMSKHMSLLADLKT(15)DKYND
13. LEGVFTPLEVGGALFAACIHDVDHPGLTNQFLVNSSSELALMYN-DESVLENHHLAVAFKLL-QNQG-CDIFCNMQKKQRQTLRKMVIDIVLSTDMSKHMSLLADLKT(15)DNYTD
14. LDAVFTDLEILAALFAAAIHDVDHPGVSNQFLINTNSELALMYN-DESVLENHHLAVGFKLL-QEDN-CDIFQNLSKRQRQSLRKMVIDMVLATDMSKHMTLLADLKT(15)DNYSD
15. LDAVFTDLEILAAIFAAAIHDVDHPGVSNQFLINTNSELALMYN-DESVLENHHLAVGFKLL-QEEH-CDIFMNLTKKQRQTLRKMVIDMVLATDMSKHMSLLADLKT(15)DNYTD
16. LEAVFTDLEILAALFASAIHDVDHPGVSNQFLINTNSELALMYN-DASVLENHHLAVGFKLL-QAEN-CDIFQNLSAKQRLSLRRMVIDMVLATDMSKHMNLLADLKT(15)DNYSD
17. LEAVFTDLEILAAIFASAIHDVDHPGVSNQFLINTNSELALMYN-DSSVLENHHLAVGFKLL-QEEN-CDIFQNLTKKQRQSLRKMVIDIVLATDMSKHMNLLADLKT(15)DNYSD
18. IQKRLTDLEILALLIAALSHDLDHRGVNNSYIQRSEHPLAQLY--CHSIMEHHHFDQCLMIL-NSPG-NQILSGLSIEEYKTTLKIIKQAILATDLALYIKRRGEFFE( 9)EDPHQ
19. LKRYFTDLEALAMVTAAFCHDIDHRGTNNLYQMKSQNPLAKLH--GSSILERHHLEFGKTLL-RDES-LNIFQNLNRRQHEHAIHMMDIAIIATDLALYFKKRTMFQK(19)LEQTR
20. LKSYYTDLEAFAMVTAGLCHDIDHRGTNNLYQMKSQNPLAKLH--GSSILERHHLEFGKFLL-SEET-LNIYQNLNRRQHEHVIHLMDIAIIATDLALYFKKRAMFQK(19)LETTR
21. IKKYYTDLEAFAMVAAAFCHDIDHRGTNNLYQMKSAAPLAKLH--GSSILERHHLEYSKTLL-QDES-LNIFQNLNKRQFETVLHLFEVAIIATDLALYFKKRTMFQK(19)IDPTK
22. LKKYYTDLEAFAMLAAAFCHDIDHRGTNNLYQMKSTSPLARLH--GSSILERHHLEYSKTLL-QDES-LNIFQNLNKRQFETVIHLFEVAIIATDLALYFKKRTMFQK(19)VDPTK
23. LANSVTPWDILLSLIAAATHDLDHPGVNQPFLIKTNHYLATLYK-NTSVLENHHWRSAVGLL-RE---SGLFSHLPLESRQQMETQIGALILATDISRQNEYLSLFRS( 9)EDTRH
24. LKDN--PVQTLLLCMAAIGHDVGHPGTNNQLLCNCESEVAQNFK-NVSILENFHRELFQQLL-SEH--WPLKLSISKKKFD----FISEAILATDMALHSQYEDRLMH( 0)ENPMK
            o     o   **  * *  o            *       ooo*  *     o                                  oo **o

 1. RSKVLCLLLHMADISHPGKEWETHQEFSSRLVEEFFRQGDQEMQ-LGRTCSPLCN( 0)RNTTS-VPECQLGFIDYIVSPSFDVCG--DLL(25)WEKPINSNKCAW  212- 538
 2. KNKALCLIVHACDISHPAKPWNLHERWTEGVLEEFFRQGDLEAS-MGLPYSPLCD( 0)RHTVH-VADSQIGFIDFIVEPTMVVCG--ELL(65)WMKFLHENKAHW  267- 627
 3. RAKTMSLILHAADISHPAKSWKLHYRWTMALMEEFFLQGDKEAE-LGLPFSPLCD( 0)RKSTM-VAQSQIGFIDFIVEPTFSLLT--DST(67)LVDIIQQNKERW  153- 515
 4. KPKALSLLLHAADISHPTKQWLVHSRWTKALMEEFFRQGDKEAE-LGLPFSPLCD( 0)RTSTL-VAQSQIGFIDFIVEPTFSVLT--DVA(44)WVKRIQENKQKW  157- 496
 5. KPKALSLMLHTADISHPAKAWDLHHRWTMSLLEEFFRQGDREAE-LGLPFSPLCD( 0)RKSTM-VAQSQVGFIDFIVEPTFTVLT--DMT(63)WTEVVHINRERW  162- 521
 6. HRLMCSLIMTSCDLTASCKTWESNKAISDLIYQEFFSQGDLEKA-LGVTPSEMMD( 0)RDRAF-IPEQQLQFLDNIAGPVYQLLS--RLL( 5)AYHTLLDNREQW  497- 803
 7. HRLLLCLLMTSCDLSDQTKGWKTTRKIAELIYKEFFSQGDLEKA-MGNRPMEMMD( 0)REKAY-IPELQISFMEHIAMPIYKLLQ--DLF( 5)LYERVASNREHW  589- 895
 8. RLTVMKMIMKMADISTPTKSYELHRAWTKLITEEFYQQGEQEIV-LGLPPTIFLD( 0)RKHPEKLPAVQLSFIQNLVSPLFHACAEAGVI(64)ILSNLQMNYEAW  549- 943
 9. RLLVCQMCIKLADINGPAKCKELHLQWTDGIVNEFYEQGDEEAS-LGLPISPFMD( 0)RSAPQ-LANLQESFISHIVGPLCNSYDSAGLM(52)ITQHLLQNHKMW  685-1086
10. RLLVCQVCIKLADINGPAKVRDLHLKWTEGIVNEFYEQGDEEAN-LGLPISPFMD( 0)RSSPQ-LAKLQESFITHIVGPLCNSYDAAGLL(51)LMHHLTENHKIW  670-1072
11. RVEVLQGLLHCADLSNQTKSFHIASQWSQRIMEESYRQGDEERK-LGIPVNPLGD( 0)RS-VS-IEKCQVTFIDYIVYPLWETWA--ELV( 6)IVCSLMDTRDQW  351- 660
12. KIQVLQSMIHLADLSNPTKPIELYQQWNQRIMEEYWRQGDKEKE-LGLEISPMCD( 0)RGNVT-IEKSQVGFIDYIVHPLYETWA--DLV( 6)ILDQLEENREWY  218- 528
13. RIQVLENLVHCADLSNPTKPLPLYKRWVALLMEEFFLQGDKERE-SGMDISPMCD( 0)RHNAT-IEKSQVGFIDYIVHPLWETWA--SLV( 6)ILDTLEENRDYY  375- 686
14. RIQVLRNMVHCADLSNPTKPLELYRQWTDRIMAEFFQQGDRERE-RGMEISPMCD( 0)KHTAS-VEKSQVGFIDYIVHPLWETWA--DLV( 6)ILDTLEDNRDWY  368- 679
15. RIQVLRNMVHCADLSNPTKSLELYRQWTDRIMEEFFQQGDKERE-RGMEISPMCD( 0)KHTAS-VEKSQVGFIDYIVHPLWETWA--DLV( 6)ILDTLEDNRNWY  341- 652
16. RIQVLQNLVHCADLSNPTKPLPLYRQWTDRIMAEFFQQGDRERE-SGLDISPMCD( 0)KHTAS-VEKSQVGFIDYIAHPLWETWA--DLV( 6)LLDTLEDNREWY  311- 622
17. RIQVLQNMVHCADLSNPTKPLQLYRQWTDRIMEEFFRQGDRERE-RGMEISPMCD( 0)KHNAS-VEKSQVGFIDYIVHPLWETWA--DLV( 6)ILDTLEDNREWY  261- 572
18. KELFLAMLMTACDLSAITKPWPIQQRIAELVATEFFDQGDRERKELNIEPADLMN( 0)REKKNKIPSMQVGFIDAICLQLYEALT--HVS( 5)LLDGCRKNRQKW  537- 843
19. KEIVMAMMMTACDLSAITKPWEVQSQVALLVAAEFWEQGDLERTVLQQNPIPMMD( 0)RNKADELPKLQVGFIDFVCTFVYKEFS--RFH( 5)MLDGITNNRKEW  494- 809
20. KEIVMAMMMTACDLSAITKPWEVQSKVALLVAAEFWEQGDLERTVLDQQPIPMMD( 0)RNKAAELPKLQVGFIDFVCTFVYKEFS--RFH( 5)MFDRLQNNRKEW  492- 807
21. KEVIMAMMMTGCDLSAITKPWEVQSKVALMVANEFWEQGDLERTVLQQQPIPMMD( 0)RNKGDELPKLQVGFIDFVCTFVYKEFS--RFH( 5)MFDGLQNNRVEW  497- 812
22. KEIIMAMMMTACDLSAITKPWEVQSQVALMVAAEFWEQGDLERTVLQQQPIPMMD( 0)RNKRDELPKLQVGFIDFVCTFVYKEFS--RFH( 5)MLSGLQNNRVEW  497- 812
23. RHLVLQMALKCADICNPCRTWELSKQWSEKVTEEFFHQGDIEKK-YHLGVSPLCD( 0)RHTES-IANIQIGFMTYLVEPLFTEWA--RFS( 7)MLGHVGLNKASW  163- 467
24. QITLISLIIKKAADISNVTRTLSISARWAYLITLEFNDCALLET--FHKAHRPEQD(33)KDHPH-IPNGQIFFINTFAEVFFNALSQ-KFS( 4)LSDNVKINKEYW  204- 519
        o    *o    o           o   *    o  *            o     o     o   *  *o                    o
```

Fig. 1. Alignments of the amino acid sequences of (a) the catalytic domain, (b) the CaM interaction domain, (c) the cGMP binding domain, and (d) the upstream conserved region 1 and 2. * and ○, amino acid positions that are occupied by identical and chemically similar amino acids for all the sequences compared, respectively. Gaps (-) were inserted to increase sequence similarity. For regions where unambiguous alignments are not possible, the numbers of amino acids involved are shown in parentheses. The start and end positions of the aligned region are also shown for each sequence. For sequence data sources and accession numbers, see Section 2.

b)

```
1. sponge EFPDE1      LRLLLKQLDVGEMPSVDALKTTISFAADVLNS--------IAKNGSVVPSDDKFQSVPDEKVREWLSQTFTREESTYTNSQQPVQRKFKSVAQVVLIGQ
2. nematode T04D3.3   LRYILHQLNSGQLP-LEDLKRNIEYAALVLETAYMDETRRICDEDDDLA-EVTPETVPDE-VREWLAATFTRQNA--GKKR--DKPKFKSVANAIRTGI
3. human PDE1A        LRCLVKQLERGDVN-VVDLKKNIEYAASVLEAVYIDETRRLLDTEDELS-DIQTDSVPSE-VRDWLASTFTRKMG-MTKKKPEEKPKFRSIVHAVQAGI
4. human PDE1B        LRYMVKQLENGEIN-IEELKKNLEYTASLLEAVYIDETRQILDTEDELQ-ELRSDAVPSE-VRDWLASTFTQQAR-AKGRRAEEKPKFRSIVHAVQAGI
5. human PDE1C        LRSLVKQLERGEAS-VVDLKKNLEYAATVLESVYIDETRRLLDTEDELS-DIQSDAVPSE-VRDWLASTFTRQMG-MMLRRSDEKPRFKSIVHAVQAGI
                      ** ooo**o *o    o   **   o oo* o*oo          o        o  o    oo** * **o**o ***                    o*o*o   o  *

1. YIDNIYHEGSSIVQ-QYPPGV  99-209
2. FFEKLFRKQ-QVVQCPIPPEI  77-188
3. FVERMYRKTYHMVGLAYPAAV  35-150
4. FVERMFRRTYTSVGPTYSTAV  39-154
5. FVERMYRRTSNMVGLSYPPAV  44-159
   o o ooo     *    oo o
```

c)

```
1. sponge EFPDE2      MLRV--CGELIDLDVVSLSIKILKHIMEVSNAKACTLFVYEDVTQEL--VAYTYNGV-----------PLDKEIRKPVSSSIYGECFTTGEIINISNVP
2. human PDE2A        ILQL--CGELYDLDASSLQLKVLQYLQQETRASRCCLLLVSEDNLQ---LSCKVIGDK----------VLGEEVSFPLT-GCLGQVVEDKKSIQLKDLT
3. bovine PDE5A       LLELV-KDISSHLDVTALCHKIFLHIHGLISADRYSLFLVCEDSSNDKFLISRLFDVAEGSTLEE---ASNNCIRLEWNKGIVGHVAAFGEPLNIKDAY
4. human PDE6A        IFDLL-RDFQENLQTEKCIFNVMKKLCFLLQADRMSLFMYRTRNGIAE-LATRLFNVHKDAVLEDCLVMPDQEIVFPLDMGIVGHVAHSKKIANVPNTE
5. human PDE6B        LLELV-QDMQESINMERVVFKVLRRLCTLLQADRCSLFMYRQRNGVAE-LATRLFSVQPDSVLEDCLVPPDSEIVFPLDIGVVGHVAQTKKMVNVEDVA
6. chicken PDE6C      LFELLTEIQDEAGSMEKIVHKTLQRLSQLLARDRCSMFICRSRNGIPE-VATRLLNVTPTSKFEDNLVNPDKETVFPLDIGIAGWVAHTKKFFNIPDVK
7. human PDE6C        CLELLWTVQEEGGTPEQGVHRALQRLAHLLQADRCSMFLCRSRNGIPE-VASRLLDVTPTSKFEDNLVGPDKEVVFPLDIGIVGWAAHTKKTHNVPDVK
                         o                  o         o o         o                                          o  *         oo o

1. QDMRFNPRIDTIKGYEPTHLLCIPVKDRGAQGSQSVIGVLVVCDKN--NDRPFTKHDEEGLLYSLHFCSSMLNNTLVYQRELALKKQNEVLLQVAKNLFTSLDNLVSLLRGIMNAAS
2. SE-DVQQ-LQSMLGCELQAMLCVPVISRATDQ---VVALACAFNKL--EGDLFTDEDEHVIQHCFHYTSTVLTSTLAFQKEQKLKCECQALLQVAKNLFTHLDDVSVLLQEIITEAR
3. EDPRFNAEVDQITGYKTQSILCMPIKNHREE----VVGVAQAINKKSGNGGTFTEKDEKDFAAYLAFCGIVLHNAQLYETSLLENKRNQVLLDLASLIFEEQQSLEVILKKIAATII
4. EDEHFCDFVDILTEYKTKNILASPIMNGKD-----VVAIIMAVNKV--DGSHFTKRDEEILLKYLNFANLIMKWYHLSYLHNCETRRGQILLWSGSKVFEELTDIERQFHKALYTVR
5. ECPHFSSFADELTDYKTKNMLATPIMNGKD-----VVAVIMAVNKL--NGPFFTSEDEDVFLKYLNFATLYLKIYHLSYLHNCETRRGQVLLWSANKVFEELTDIERQFHKAFYTVR
6. KNNHFSDYLDKKTGYTTVNMMAIPITQGKE-----VLAVVMALNKL--NASEFSKEDEEVFKKYLNFISLVLRNHHTSYLYNIESRRSQMLLWSANKVFEELTDIERQFHKALYTIR
7. KNSHFSDFMDKQTGYVTKNLLATPIVVGKE-----VLAVIMAVNKV--NASEFSKQDEEVFSKYLNFVSIILRLHHTSYMYNIESRRSQILMWSANKVFEELTDVERQFHKALYTVR
         o          oo  *o          *ooo    o*    o    *o  **          o     o                       o *o  o   o*      o

1. SLTNAERCSLFLLDKSRSNLVATVFNGDVLKER-----------------------------------TLTIKVGQGIAGYVAKTGTIVNIVDAQKHPQFF---AEVDKSTGFHTK
2. NLSNAEICSVFLLD--QNELVAKVFDGGVVDDESY---------------------------------EIRIPADQGIAGHVATTGQILNIPDAYAHPLFY---RGVDDSTGFRTR
3. SFMQVQKCTIFIVDEDCSDSFSSVFHMECEELEKSSDTLTRE---------------------------RDANRINYMYAQYVKNTMEPLNIPDVSKDKRFPWTNENMGNINQQCIR
4. AFLNCDRYSVGLLDMTKQKEFFDVWPVLMGEVPPYSGPRTPDGREINFYKVIDYILHGKEDIKVIPNPPPDHWALVSGLPTYVAQNGLICNIMNAPSEDFFAFQKEPLDES-GWMIK
5. AYLNCERYSVGLLDMTKEKEFFDVWSVLMGESQPYSGPRTPDGREIVFYKVIDYILHGKEEIKVIPTPSADHWALASGLPSYVAESGFICNIMNASADEMFKFQEGALDDS-GWLIK
6. MYLNCERYSVGLLDMTKEKEFYDEWPIRLGEAEPYKGPKTPDGREVNFYKIIDYILHGKEEIKVIPTPPADHWCLISGLPTYVAENGFICNMMNAPADEYFTFQKGPVDET-GWVIK
7. TYLNCERYSIGLLDMTKEKEFYDEWPIKLGEVEPYKGPKTPDGREVNFYKIIDYILHGKEEIKVIPTPPADHWTLISGLPTYVAENGFICNMMNAPADEYFTFQKGPVDET-GWVIK
     o o  oo oo*         o                                                 o  *      *o o     *     o     o

1. HILCFPIMD-NNGVVVGVAELCNKING-----KFFTKYDEELARTFSAYCGISIYHSKLYETVMASQGRSSLATELMLYHMKIRPDE  132-489
2. NILCFPIKN-ENQEVIGVAELVNKING-----PWFSKFDEDLATAFSIYCGISIAHSLLYKKVNEAQYRSHLANEMMMYHMKVSDDE  230-581
3. SLLCTPIKNGKKNKVIGVCQLVNKMEETTGKVKAFNRNDEQFLEAFVIFCGLGIQNTQMYEAVERAMAKQMVTLEVLSYHASAAEEE  142-526
4. NVLSMPIVN-KKEEIVGVATFYNRKDG-----KPFDEMDETLMESLTQFLGWSVLNPDTYESMNKLENRKDIFQDIVKYHVKCDNEE   61-464
5. NVLSMPIVN-KKEEIVGVATFYNRKDG-----KPFDEQDEVLMESLTQFLGWSVMNTDTYDKMNKLENRKDIAQDMVLYHVKCDRDE   59-462
6. NVLSLPIVN-KKEEIVGVATFYNRKDG-----KPFDEYDEQIIETLTQFLGWSVLNTDTYDKMNKLENRKDIAQEMLMYQTKATPTE   62-466
7. NVLSLPIVN-KKEDIVGVATFYNRKDG-----KPFDEHDEYITETLTQFLGWSLLNTDTYDKMNKLENRKDIAQEMLMNQTKATPEE   62-466
    o*  ** o     oo**   *o o      *   **    o   o *oo  o  *  o    o  o ooo        *
```

d)

```
1. sponge EFPDE4      PFRRRSSFLHQSDAEEP-SSKSLSRVSSCGS-------------HVGDETFSTPFAQILANFRCVRANLSTLLDEEKLTTV----------------
2. nematode R153.1    --------SISLINNNNRSVRRKFGESGFTL-------------RHGDDLIVTPFAQLLASLRNVRSNLISITNIQNSDDSRHANRSAKRPP------
3. Drosophila dunce   PQRRE-SFLYRSDSDFEMSPKSMSRNSSIASERFKEQEASILVDRSHGEDLIVTPFAQILASLRSVRNNLLSLTNVPASNKSRRPNQSSSASRSGNPPG
4. human PDE4A        SQRRE-SFLYRSDSDYDMSPKTMSRNSSVTSE------------AHAEDLIVTPFAQVLASLRSVR------SNFSLLTNV--PVPSNKRSPLGGPTP
5. human PDE4B        SQRRE-SFLYRSDSDYDLSPKAMSRNSSLPSE------------QHGDDLIVTPFAQVLASLRSVR------NNFTILTNL--HGTSNKRSPAASQPP
6. human PDE4C        SQRRE-SFLYRSDSDYELSPKAMSRNSSVASD------------LHGEDMIVTPFAQVLASLRTVR------SNVAALARQQ-CLGAAKQGPVGNPSS
7. human PDE4D        SQRRE-SFLYRSDSDYDLSPKSMSRNSSIASD------------IHGDDLIVTPFAQVLASLRTVR------NNFAALTNSQ-DRAPSKRSPTCNQPS
                                        o  * o    *o o                ooo  *****o**  * **      o

1. -SEDTTVDERSAQLAKDTMMEFDWCMQKLETVDRVK-MGSLAQDKFQRVMSRELTQLSERSLSGHRVAEWVQDITRSDRDDEIDRA  148-298
2. -LHNIELPDDVVHCAHDTLEELDWCLDQLETIQTHRSVSEMASSKFRKMLNKELSHFAESSKSGTQVSKFLITTYMDKEEDEPSIE    1-157
3. -APLSQGEEAYTRLATDTIEELDWCLDQLETIQTHRSVSDMASLKFKRMLNKELSHFSESSRSGNQISEYICSTFLDKQQEFDLPS  121-303
4. VCKATLSEETCQQLARETLEELDWCLEQLETMQTYRSVSEMASHKFKRMLNRELTHLSEMSRSGNQVSEYISTTFLDKQNEVEIPS  140-302
5. VSRVNPQEESYQKLAMETLEELDWCLDQLETIQTYRSVSEMASNKFKRMLNRELTHLSEMSRSGNQVSEYISNTFLDKQNDVEIPS  128-290
6. SNQLPPAEDTGQKLALETLDELDWCLDQLETLQTRHSVGEMASNKFKRILNRELTHLSETSRSGNQVSEYISRTFLDQQTEVELPK  102-265
7. INKATITEEAYQKLASETLEELDWCLDQLETLQTRHSVSEMASNKFKRMLNRELTHLSEMSRSGNQVSEFISNTFLDKQHEVEIPS   49-212
          o     * o*o * ***oo ***oo  o oo o*   ** ooo o**o  o* * **  oo oo          o
```

Fig. 1 (*continued*).

three independent clones were isolated for sponge PDE genes and sequenced by the dideoxy chain termination method [8] using synthetic oligonucleotide as primers. The full-length sponge PDE coding sequences were obtained by 5′ and 3′ rapid amplification of cDNA ends (Gibco BRL) [9]. *E. fluviatilis* genomic DNA fragments containing the PDE sequences were identified by Southern blot analysis with specific probes.

2.2. Sequence data

Accession numbers of sequence data from GenBank release 101.0 and PIR(*) database release 51.0 are as follows (#, this work): sponge EFPDE1 (AB017021#); *Caenorhabditis* T04D3.3 (Z81114); human PDE1A (U40370); human PDE1B (U56976); human PDE1C (U40371); sponge EFPDE2 (AB017022#); human PDE2A (U67733); sponge EFPDE3 (AB017023#); human PDE3A

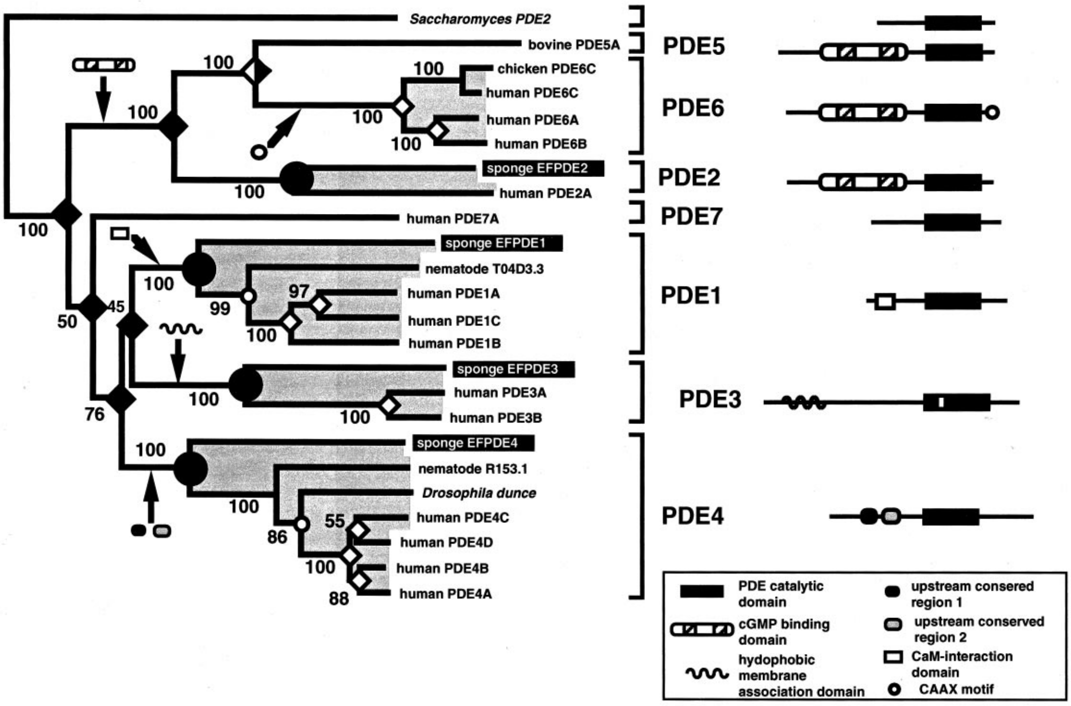

Fig. 2. Phylogenetic tree of the PDE family inferred from the catalytic domain. The tree was inferred by the NJ method using a fungal PDE as outgroup. The number at each branch point represents the bootstrap probability that two lineages join together to form a cluster. Filled circle, parazoan-eumetazoan split; open circles, human-*Drosophila* (or nematode) split; filled rhombi, gene duplications that gave rise to different subtypes; half filled rhombus, gene duplications whose divergence time is unknown; open rhombi, gene duplications in the same subtype. The branch length is proportional to the number of accumulated amino acid substitutions. Domains that were integrated during evolution by domain shuffling are shown. Data on the domain structure were taken from published papers [18–20].

(U36798); human PDE3B (X95520); sponge EFPDE4 (AB017024#); *Caenorhabditis* R153.1 (U28729); *Drosophila dunce* (S65543*); human PDE4A (L20965); human PDE4B (L20966); human PDE4C (U88713); human PDE4D (U50159); bovine PDE5A (L16545); human PDE6A (M26061); human PDE6B (X66142); chicken PDE6C (L29233); human PDE6C (X94354); human PDE7A (L12052); *Saccharomyces* PDE2 (M14563).

2.3. Alignment and phylogenetic tree inference

Optimal alignment of sequences was obtained by the methods of Needleman and Wunsch [10] and Berger and Munson [11], together with manual inspections. The number k_{aa} of amino acid substitutions per site or evolutionary distance was calculated by the method of Jukes and Cantor [12] as $k_{aa} = -\ln(1-K_{aa})$ for regions where unambiguous alignment is possible, where K_{aa} represents the amino acid difference per residue between sequences compared; amino acid sites where gaps exist in the alignment were excluded from the calculation. The evolutionary distance was applied to phylogenetic inference by neighbor-joining (NJ) method [13]. Bootstrap analysis was carried out by the method of Felsenstein [14]. The phylogenetic tree of the PDE family inferred by the NJ method was reexamined by the maximum likelihood (ML) method of protein phylogeny [15,16] based on the JTT model (PROTML version 2.2 in Adachi and Hasegawa's program package MOLPHY).

3. Results and discussion

To determine whether multiple PDEs exist in the sponge lacking the cell cohesiveness and coordination typical of eumetazoans [17], and whether the domain organization of the sponge PDEs is similar to that of vertebrates, we carried out cloning of PDE cDNAs from the freshwater sponge *E. fluviatilis* by RT-PCR. Four cDNAs, EFPDE1, EFPDE2, EFPDE3, and EFPDE4, were obtained, which are closely related in amino acid sequence to members of the vertebrate PDE1, PDE2, PDE3, and PDE4 subtypes over the entire regions, respectively. Including the four sponge cDNAs, alignment of the catalytic domain sequences of the PDE family members is shown in Fig. 1a. Alignments of the calmodulin (CaM) interaction domain of the PDE1 subtype, the cGMP

Table 1
Comparison of the evolutionary rates of PDE subtypes between the first and later periods of animal evolution

Subtype	First period v_I	Later period v_{II}	v_I/v_{II}
PDE1	1.6×10^{-9}	0.39×10^{-9}	4.1
PDE2	2.9	0.33	8.8
PDE3	2.2	0.35	6.3
PDE4	1.5	0.38	3.9
Mean	2.1	0.36	5.8

The whole animal lineage from the common ancestor of animals and fungi (or plants) to extant animals was divided into two periods, the first period (I) and the later period (II), by tentatively defining the divergence time of parazoans and eumetazoans as the boundary. The evolutionary rates v_I/site/year in the first period and v_{II} in the later period were calculated as follows: for each subtype, the number of amino acid substitutions per site (k_{aa}) accumulated in each of the first and later periods was calculated based on the branch lengths of the trees in Fig. 2; the k_{aa} of the first period was calculated from the branch length between the deepest gene duplication in the tree and the deepest node in each cluster corresponding to the subtype, and for the k_{aa} of the later period, the average length of different branches between the deepest node of the subtype and the extant species was used. This method gives an underestimate of k_{aa} in the first period and an overestimate in the later period. Using the k_{aa} values and assuming divergence times of 1070 Myrs ago and 940 Myrs ago for the animal-fungus (or plant) split and the parazoan-eumetazoan split [21], respectively, the evolutionary rates v_I and v_{II} were calculated.

binding domain of PDE2, PDE5 and PDE6, and the upstream conserved regions of PDE4 [18] are shown in Fig. 1b–d. The vertebrate PDE3 has a stretch of hydrophobic amino acids (membrane association domain) near the N-terminal end [19]. The hydrophobic region is also found in the sponge PDE3 at the equivalent position. As the alignments of Fig. 1 show, four sponge cDNAs contain domains similar to the respective domains of vertebrate subtypes in the corresponding positions; the cGMP binding domain of EFPDE2 is more closely related in amino acid sequence to that of vertebrate PDE2 (43% identical) than those of vertebrate PDE5 (29%) and PDE6 (24–26%). Thus at least in four subtypes, the domain organization is virtually identical between vertebrates and sponge.

Using a fungal PDE as an outgroup, a phylogenetic tree was inferred by the NJ method [13] based on the alignment of the catalytic domain sequences (Fig. 2). As Fig. 2 shows, the sponge EFPDE1, EFPDE2, EFPDE3, and EFPDE4 belong to PDE1, PDE2, PDE3, and PDE4 subtypes, respectively. This result, together with the similarity of domain organization, strongly suggests that the isolated sponge cDNAs are homologs of vertebrate PDEs.

The phylogenetic tree of the PDE family including sponge homologs provides clear-cut evidence for subtype duplication and domain shuffling in the early evolution of animals. Five subtype duplications out of six are very old, going back to dates before the parazoan-eumetazoan split about 940 Myrs ago [21], the earliest branching among extant animal phyla; the date of subtype duplication that gave rise to PDE5 and PDE6 is unknown. The same result was also obtained from the phylogenetic tree inferred by the ML method [15,16]. In addition, the number of subtype duplication was reexamined statistically by 1000 bootstrap resamplings, as described previously [5], and we obtained average numbers of subtype duplications before and after the sponge-eumetazoan split of 5.0 ± 0.1 and 0.0 ± 0.0, respectively.

Because *Saccharomyces cerevisiae* has a single copy of PDE which shares similarity with animal PDEs, it is highly likely that all seven animal PDE subtypes diverged from a common ancestral gene by gene duplication in the early evolution of animals, and most, if not all, of the subtypes were established within a period ('first period') between animal-fungus-plant splits about 1070 Myrs ago [21] and the parazoan-eumetazoan split about 940 Myrs ago. In contrast, over the long evolutionary time span of animal evolution since the parazoan-eumetazoan split to the present time ('later period'), no subtype duplication has been observed. Because the sponge EFPDE1, EFPDE2, EFPDE3, and EFPDE4 share domain organization with the respective vertebrate PDEs, it is possible to infer from the phylogenetic tree that five different domains out of six excluding the catalytic domain were integrated into ancestral genes in the first period by domain shuffling. Assuming the simplest structure in the common ancestor and parsimonious domain shuffling, it is possible to trace the evolution of domain organization (Fig. 2).

In addition, the branch length of the tree shows a rapid accumulation of amino acid substitutions in the first period. To obtain a qualitative estimate, the evolutionary rates v_{I} of amino acid substitutions of each subtype in the first period and v_{II} in the later period were calculated based on the branch lengths of the tree, assuming the divergence times of the parazoan-eumetazoan split and animal-fungus-plant splits to be 940 Myrs ago and 1070 Myrs ago, respectively [21]. The result is summarized in Table 1. The v_{I} is remarkably high, being 5.8 times higher than v_{II} on the average of four subtypes. The explosive subtype duplication and the rapid evolutionary rate in the first period are also found in other gene families involved in signal transduction and developmental control, including the Gα, PTK (Suga et al., submitted), phospholipase C, protein kinase C (Koyanagi et al., in preparation), protein tyrosine phosphatase (Ono et al., in preparation) and *Pax* families [22].

In the PDE1, PDE3, PDE4 and PDE6 subtypes, further gene duplications are observed, by which multiple isoforms were created in each subtype. As the PDE4 subtype shows, the isoform duplication postdates the vertebrate-arthropod split, which is consistent with our previous analyses for many gene families involved in signal transduction [1,5]. In addition, previous analyses showed that most isoform duplications were completed before the fish-tetrapod split [1,5].

In summary, the majority of the present-day PDE subtypes with distinct domain organization and primary function were created from an ancestral gene by explosive gene duplication and domain shuffling in the early evolution of animals, accompanying rapid amino acid substitutions, and were established at ancient dates before the parazoan-eumetazoan split. After the separation from arthropods, isoform duplications frequently occurred in chordate lineages, possibly in the first half of chordate evolution [1,5]. Thus the PDE family, as well as other gene families involved in signal transduction and developmental control, increased the multiplicity of family members intermittently, but not gradually during animal evolution. The extensive subtype duplication in the first period may be related to the evolution of multicellularity. Unexpectedly, the frequency of gene duplication is extremely low at the Precambrian-Cambrian boundary when the Cambrian explosion occurred, and thus it is reasonable to consider that there is no direct link between the burst of gene duplication and the Cambrian explosion. It seems conceivable that animals underwent the Cambrian explosion by using already existing genes, and not by creating new genes with novel functions. Examinations based on factors other than gene duplication and domain shuffling are necessary for understanding the molecular mechanism of the Cambrian explosion.

Acknowledgements: We thank Prof. Y. Watanabe for kindly providing us the freshwater sponge (*Ephydatia fluviatilis*). This work was supported in part by grants from the Ministry of Education, Science, Sports and Culture of Japan.

References

[1] Iwabe, N., Kuma, K. and Miyata, T. (1996) Mol. Biol. Evol. 13, 483–493.
[2] Beavo, J.A., Conti, M. and Heaslip, R.J. (1994) Mol. Pharmacol. 46, 399–405.
[3] Houslay, M.D. and Milligan, G. (1997) Trends Biochem. Sci. 22, 217–224.
[4] Sass, P., Field, J., Nikawa, J., Toda, T. and Wigler, M. (1986) Proc. Natl. Acad. Sci. USA 83, 9303–9307.
[5] Suga, H., Kuma, K., Iwabe, N., Nikoh, N., Ono, K., Koyanagi, M., Hoshiyama, D. and Miyata, T. (1997) FEBS Lett. 412, 540–546.
[6] Conway Morris, S. (1993) Nature 361, 219–225.
[7] Seimiya, M., Ishiguro, H., Miura, K., Watanabe, Y. and Kurosawa, Y. (1994) Eur. J. Biochem. 221, 219–225.

[8] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[9] Frohman, M.A., Dush, M.K. and Martin, G.R. (1988) Proc. Natl. Acad. Sci. USA 85, 8998–9002.
[10] Needleman, S.B. and Wunsch, C.D. (1970) J. Mol. Biol. 48, 443–453.
[11] Berger, M.P. and Munson, P.J. (1991) CABIOS 7, 479–484.
[12] Jukes, T.H. and Cantor, C.R. (1969) in: Mammalian Protein Metabolism III (Munro, H.N., Ed.), pp. 21–132, Academic Press, New York.
[13] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.
[14] Felsenstein, J. (1985) Evolution 39, 783–791.
[15] Kishino, H., Miyata, T. and Hasegawa, M. (1990) J. Mol. Evol. 30, 151–160.
[16] Adachi, J. and Hasegawa, M. (1992) in: Computer Science Monographs, No. 27, The Institute of Statistical Mathematics, Tokyo.
[17] Margulis, L. and Schwartz, K.V. (1988) Five Kingdoms, 2nd edn., W.H. Freeman and Company, New York.
[18] Bolger, G.B. (1994) Cell. Signal. 6, 851–859.
[19] Taira, M., Hockman, S.C., Calvo, J.C., Taira, M., Belfrage, P. and Manganiello, V.C. (1993) J. Biol. Chem. 268, 18573–18579.
[20] Li, T., Volpp, K. and Applebury, M.L. (1990) Proc. Natl. Acad. Sci. USA 87, 293–297.
[21] Nikoh, N., Iwabe, N., Kuma, K., Ohno, M., Sugiyama, T., Watanabe, Y., Yasui, K., Shicui, Z., Hori, K., Shimura, Y. and Miyata, T. (1997) J. Mol. Evol. 45, 97–106.
[22] Hoshiyama, D., Suga, H., Iwabe, N., Koyanagi, M., Nikoh, N., Kuma, K., Matsuda, F., Honjo, T. and Miyata, T. (1998) J. Mol. Evol. (in press).